Robust Harris Detector Corresponding and Calculates the Projection Error Using the Modification of the Weighting Function

A. Chater and A. Lasfar

Abstract—**Epipolar geometry is a key point in computer vision and the fundamental matrix estimation. The relational view can be obtained from the fundamental matrix. In this way, we interest to calculate an exact matrix based from characteristics unequally distributed in complex scene images. This paper presents a method based on the detection of points by the Harris detector, after we develop a new modification of the multi-level function related to the M-estimator algorithm. The experimental comparisons were conducted by a simulation between RANSAC, LMeds, and M-Estimator in order to estimate the projection error. As a result, the proposed method gives a significant improvement and performance with a low projection error compared to other methods.**

Index Terms—**Epipolar geometry, fundamental matrix, weighting function, projection error, Harris detector.**

I. Introduction

Epipolar geometry makes it possible to establish a geometrical relationship between two images taken from two different angles of view [1]. The epipolar geometry is mainly used in projective geometry, photogrammetry, and artificial vision and the main purpose of his application is the analysis of correspondences between two different views of the same image [2].

Many methods for estimating the fundamental matrix exist and almost all use only image point matches as data. These methods are distinguished by the parameterization of 'F' as well as by the optimization method. The estimation of the fundamental matrix is based on the knowledge of a certain number of pairs of points in correspondence.

Recently, many researchers have studied a several methods for estimating the fundamental matrix. For instance we found Zhang [3], Xavier and Joaquim [4] and Hartley [5], who they are interested in this area, these methods can be classified into linear and iterative or robust methods.

The results of the previous works of the linear method have shown two major defects related to the absence of rank constraint of the fundamental matrix and to the absence of normalization. To defeat this defect, non-linear methods are proposed to improve the accuracy of estimation of the fundamental matrix by minimizing the distance between the points and their epipolar lines [6], [7] , moreover to reduce the effect of potential aberrant values and to have greater

tolerance in data noise [6], [8]-[10].

The first technique M-Estimator [11] leads to a good result in the presence of noise Gaussian at the selected points of the image, but it is limited in the ability to take outliers into account. Two other techniques are classified as robust methods and they are similar: the Random Sampling Consensus (RANSAC) method [12], and the median least squares method (LMeds) [6]. These two disadvantages consist in randomly selecting all the points used for the approximation of the fundamental matrix. The experiments have shown that the LMeds technique gives a better result than the RANSAC method in terms of accuracy [13].

In this paper, we develop a robust and efficient method that depends on the correspondence of characteristics to estimate the fundamental matrix. We detonate the points of interest by using the Harris method [14]-[19]. After detecting matched characteristics, we calculate the fundamental matrix directly using both RANSAC and M-Estimator methods. The purpose of calculating the 'F' is to determine the projection error. Our work is projected on the modification of the weighting function at the level of the M-Estimator method. This new proposition makes it possible not only to estimate the matrix F but also to precise and identify the different aberrant level values from the set of point correspondences. The simulations results of the real images show that the proposed method is more robust than the other methods in terms of the level of the projection error.

II. Feature Extraction

A. Harris Corner Detection

Harris detector is a method that is based on a function of auto-correlation of the signal, that is to say on the changes of the signal in several directions [14], [15]. The Fig. 1 below shows the different steps of the Harris detector.

$$E(x, y) = \sum_{u,v} w(u,v) \left| I(x+u, y+v) - I(u,v) \right|^2 \quad (1)$$

That we can develop with Taylor's formula for writing:

$$E(x, y) = \sum_{u,v} w(u,v) \left[x \frac{dI}{dx} + y \frac{dI}{dy} + o(x^2, y^2) \right] \quad (2)$$

For small displacements we can neglect the term $o(x^2, y^2)$ from where

$$E(x, y) = Ax^2 + 2Cxy + By^2 \qquad (3)$$

With, $B = \dfrac{dI}{dy}^2 * w$ and $C = \dfrac{dI}{dx}\dfrac{dI}{dy} * w$

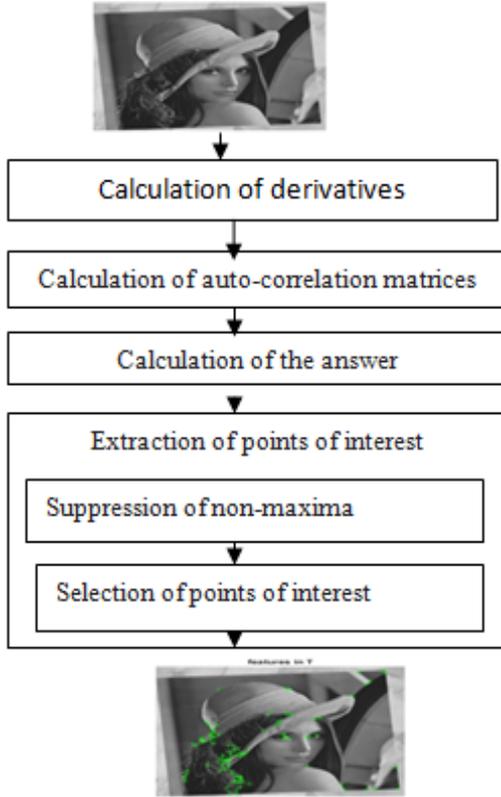In order to reduce the noise in the window used, Harris detector proposes the use of a Gaussian window:

Fig. 1. Detection of points of interest by the Harris detector method.

Harris Detector to find points of interest, the Harris detector calculated for each pixel, the auto-correlation matrix from the two components of the gradient vectors of this image. Then, the response matrix of the detector is obtained from these matrices. Finally, the points of interest, marked above with a green cross, are located on the basis of this answer.

Finally to take into account the general behavior of the function E locally, one can write:

$$E(x, y) = (x, y)M(x, y)^T \qquad (4)$$

$$with \quad M = \begin{bmatrix} A & C \\ C & B \end{bmatrix}$$

The matrix M characterizes the local behavior of the function $E(x, y)$ and. The eigenvalues of this matrix correspond in fact to the main curvatures associated wit $E(x, y)$. Ongoing back to the three case. If both curves are low eigenvalues, then the region under consideration has an approximately constant intensity. If one of the curvatures is of very strong eigenvalues so while the other is low then the region contains an outline. And if both curvatures are strong eigenvalues then the intensity varies greatly in all directions, which characterizes a wedge.

Harris offers a metric to detect corners and edges at the same time

$$R = Det(M) - K(trace^2(M)) \qquad (5)$$

As. $Det(M) = AB - C^2$ and $Trace(M) = A + B$

$$Det(M) = \lambda_1\lambda_2 \qquad and \qquad Trace(M) = \lambda_1 + \lambda_2$$

Consider the following inspired formulation for the corner response

$$R = \lambda_1\lambda_2 - k(\lambda_1 + \lambda_2)^2 \qquad (6)$$

K: Harris parameter. 0.04<k< 0.25 According to several authors, the experiment shows that an optimal number of points is obtained for a value of the order of $k = 0.04$

Knowing that if λ1 and λ2 are the eigenvalues of the matrix M. R is positive in the corner region, negative in the edge regions, and small in Hit flat region. From the new answer R we do window scanning to extract local maxima.

III. EPIPOLAR GEOMETRY AND FUNDAMENTAL MATRIX

A. Epipolar Geometry

The epipolar geometry defines the set of relations existing between two different image views of the same scene captured using a camera. The Fig. (2) presents a stereoscopic system composed of two image sensors arranged at positions o_l and o_r in the scene. A point M of space is projected respectively into P' and P'' in both left and right images. The point M and the two optical centers form a plane π, called the epipolar plane, which intersects the two image planes according to two straight lines I_l and I_r called epipolar lines. These lines cut the baseline o_l ol and o_l, a constant of the device if it is assumed that the arrangement of the two sensors is rigid, respectively at two points e_l and e_r, called epipoles. These two points are therefore also constants of the stereoscopic system.
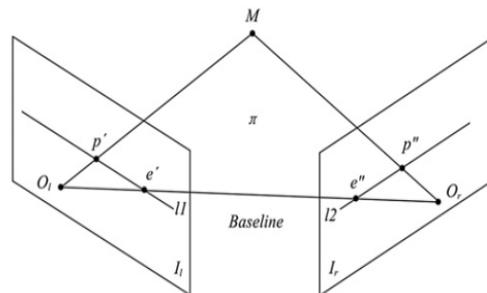

Fig. 2. Epipolar geometry of stereo images.

The relationship between the points (o_r, o_l, M, P', P''), which is expressed from two left and right images to calculate the fundamental matrix of rank equal to 2 and the

points P'', P', Equation.

$$P'^T F P'' = 0 \qquad (7)$$

B. Fundamental Matrix

The estimation of the fundamental matrix relies on the knowledge of a certain number of corresponding pairs of points. This estimation can be represented as the following equation (7).

Where 'F' is a matrix of dimension 3x3 and of rank-2, and determined from 'F' and zero, the equation is the relationship which relies the points of the left image noted $P'^T = (u_{li}, \hat{v}_{li} \ 1)$ and points of the right image noted. $P'' = (u_{ri}, \hat{v}_{ri} \ 1)^T$. This equation can be rewritten in the following linear form:

$$Af = 0 \qquad (8)$$

With:

$$A = \begin{pmatrix} u_{1l}u_{1r} & u_{1l}v_{1r} & u_{1l} & v_{1l}u_{1r} & v_{1l}v_{1r} & v_{1l} & u_{1r} & v_{1r} & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ u_{Nl}u_{Nr} & u_{Nl}v_{Nr} & u_{Nl} & v_{nl}u_{1r} & v_{1l}v_{1r} & v_{Nl} & u_{Nr} & v_{Nr} & 1 \end{pmatrix}$$

And $f = [f_1, f_2, f_3, f_4, f_5, f_6, f_7, f_8, f_9]^T$

It can be noted from the decomposition of the equation (8) that there are 9 unknown and 7 independent parameters. But it is still possible to estimate this matrix from only 7 pairs of points [6], [20]. The major advantage of this method consists on his simplicity and speed. However, the quality of the results deteriorates rapidly when a few points are poorly localized. Moreover, the solution is not always unique and the result depends on the choice of the 7 points detected by Harris detector in the set of available correspondences. This approach has been improved [16]. This author has proposed also a more robust algorithm called eight standardized point. This approach greatly improves the result of the seven-point method. In this work, we took the last algorithm of Hartely (eight standardized point).

The equation (8) is a starting point for most methods of determining the fundamental matrix. Which can be solved for up to a scale factor if N = 8, and if N is greater than that, it will be solved uniquely in a way that minimizes equation (8). In general, to solve for N equations, the singular value decomposition (SVD) of A is taken so that [21].

C. Estimation of the Fundamental Matrix and Multilevel Weighting Function

We present three nonlinear methods: RANSAC, LMeds, and M-Estimator. The first method, for its part, calculates for each value of 'F' the number of points that may be suitable (inliers). The matrix 'F' chosen is that which maximizes this number. Once the aberrant points are eliminated, the matrix 'F' is recalculated to obtain a better estimate. The LMeds method calculates for each estimation of 'F' the Euclidean distance between the points and the epipolar lines, and the choice of 'F' corresponds to the minimization of this distance.

Although, M-Estimatoris inspired by the two preceding methods, it consists in dividing the detected points into two sets: inilers and quasi-inliers. The latter method is based on solving the following expression:

$$min_F \sum_i w_i r_i^2 \qquad (9)$$

w_i: Is the weighting function.

M-estimator considers the residual of each point on the epipolar line and affects it for each outlier. Suppose that the r_i is the residual of Where

$$F = \begin{bmatrix} f_1 & f_2 & f_3 \\ f_4 & f_5 & f_6 \\ f_7 & f_8 & f_9 \end{bmatrix}$$

$$r_i = (u_{ri}, \hat{v}_{ri} \ 1) \begin{bmatrix} f_1 & f_2 & f_3 \\ f_4 & f_5 & f_6 \\ f_7 & f_8 & f_9 \end{bmatrix} (u_{li}, \hat{v}_{li} \ 1)^T$$

$$r_i = f_1 u_{il} u_{ir} + f_2 u_{il} v_{ir} + f_3 u_{il} + f_4 v_{il} u_{ir} \\ + f_5 v_{il} v_{ir} + f_6 v_{il} + f_7 u_{ir} + f_8 v_{ir} + f_9 \qquad (10)$$

Huber [22]. Has proposed the following expression for w_i

$$w_i = w_i(P_i', P_i'') = \begin{cases} 1 & |r_i| \le \sigma \\ \dfrac{\sigma}{|r_i|} & \sigma \le |r_i| < 3\sigma \\ 0 & 3\sigma < |r_i| \end{cases} \qquad (11)$$

According to [6] the robust standard deviation σ can be expressed as:

$$\sigma = \frac{median(r_i)}{\lambda}$$

Experiments have shown that the technique LMed gives better result than RANSAC method in terms of accuracy [13]. LMeds and RANSAC are considered similar, they consist to select randomly the set of points used for the approximation of the fundamental matrix. The difference exist between this two methods in the way use to determinate the chosen 'F'. LMeds calculate the 'F' from the distance between the points and the epipolar lines where it seeks to minimize the median. RANSAC calculate the matrix 'F' from the number of inliers. However, M-Estimator leads to a good result in the presence of a Gaussian noise at the selected points of the image, the robustness of this method is manifested in the reduction of aberrant values.

D. Modification of the Multi-level Weighting Function

In the literature [23], it exists two conditions to minimize the equation (9), one is detailed in [21] and another formula will be used to evaluate Sampson error, it should be noted that this error is calculated using an accurate matrix 'F'.

$$E_{sampson} = \sum_{i}^{n} w_i \frac{(P'^T F P'')^2}{(FP_i')^2_1 + (FP_i')^2_2 + (F^T P_i'')^2_1 + (F^T P_i'')^2_2}$$

(12)

where:

w_i : weighting function

$(FP_i')^2_j$, J=1, 2 the square of the jth entry for the vector (FP_i') According to Liqiang Wang and referring to the documentation [16], we choose to use the weighting function which makes it possible to calculate the weight of each point, we clarify that this function aims to improve the precision of the fundamental matrix.

The weighting function is given in equation (13) [24], [25].

$$w_i = w_i\left(P_i', P_i''\right) = \begin{cases} 1 & r \le \phi_i \sigma \\ \theta & \phi_i \sigma < r_i < \sigma \\ \dfrac{\theta \sigma}{|r_i|} & \sigma \le r_i < \varphi \sigma \\ 0 & \varphi \sigma \gtrless r \end{cases}$$

(13)

ϕ Factor to ensure the boundary of the inliers [25], and quasi-inliers; $\sigma = \dfrac{median(r_i)}{\lambda}$ Scale of the error and

$$\phi = \frac{inlier}{\text{total number of points detected}}$$

θ : Proportional factor whose scope is (0, 1); φ and λ Constants Torr [11].

This new weighting function (12) clearly shows that the weight of the corresponding points must be updated in four cases. Regarding three sets of dense cloud points: inliers for $r \le \phi_i \sigma$, quasi-inliers for $\phi_i \sigma < \sigma$ and outliers for $\sigma \le r_i < \varphi \sigma$.

In our case study, we found 70 inliers in the set of correspondences of 300 points when the fundamental matrix was estimated for the first time. So, from the next iteration, the limiting factor has become $\phi = 0.23$.

IV. RESULTS AND DISCUSSIONS

In this section, experiments were performed on two real images (Fig. 3) from two different positions by the same camera. First, we apply the Harris detector to detect the characteristic point see (Fig. 4). Then, we identify the correspondences between the images of the scene (Fig. 5), whose goal is to determine the fundamental matrix and

calculate the projection error.

Fig. 3. Original left and right stereo images of 350x400 pixels.

Fig. 4.Harris stereo images detector left and right.

Fig. 5. The points of correspondence.

In this section, we present some results of our experimentations. We compare the methods: RANSAC, LMeds and M-Estimator from the simulation data of the real images when we vary the Gaussian noise from 0 to 1. We can define a vector (a, b) which means that the random noise added at the points of correspondence are a Gaussian distribution N (0, a).

We apply this case study in our experimentations to test the robustness of our proposed method and the others two methods. The graph below shows the simulation results.
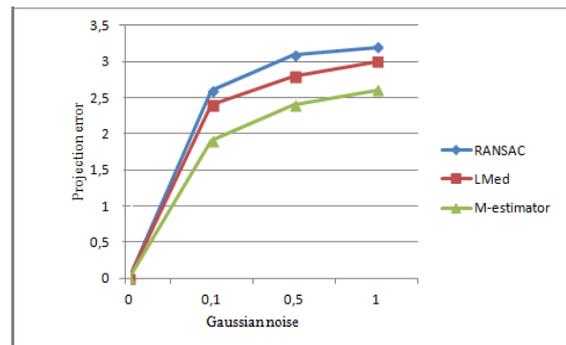

Fig. 6. Comparison of the proposed method projection error as a function of Gaussian noise.

V. CONCLUSIONS

In this article, we present a calculation of the fundamental matrix using the weighting function to improve the precision of this matrix for the interest of reducing the projection error. Then, we compare the proposed modification with the methods that already exist in the literature under the name of RANSAC method and traditional LMeds. Our modification of the weighting function is based on the distribution of the pairs of points detected by the Harris detector method in different inlier, quasi-inlier and outlier sets.

We have two added values in this work: in the first place, eliminate the descriptors of low contrast and bad correspondence correlated to the epipolar line. Secondly, insert a weighting function to plan the outliers increases the clarity of the fundamental matrix and thus increases the calculation of the projection error.

The experimental results on all the simulation data applied to the real images show that our proposed modification gives a better performance for the estimation of the fundamental matrix and the calculation of the projection error compared to the other methods

REFERENCES

[1] H. C. Longuet-Higgins, *A Computer Algorithm for Reconstructing a Scene from Two Projections*, Morgan Kaufmann Publishers Inc., 1987.

[2] S. Chaudhury *et al*., "Multiple view 3-D reconstruction in water," in *Proc. 2015 Fifth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG)*, pp. 1-4, IEEE, 2015.

[3] Z. Zhang, "Determing the epipolar geometry and its uncertainty: A review," *International Journal of Computer Vision*, vol. 27, no. 2, pp. 161–198, 1996.

[4] A. Xavier and S. Joaquim, "Overall view regarding fundamental matrix estimation," *Image and Vision Computing*, vol. 21, no. 2, pp. 205–220, 2003.

[5] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, UK, 2000.

[6] Z. Zhang, "Determining the epipolar geometry and its uncertainty: A review," *Int. J. Comput. Vis*., vol. 27, no. 2, pp. 161–195, 1998.

[7] R. Hartley and A. Zisserman, "Multiple view geometry in computer vision, 2004.

[8] A. Bartoli and P. Sturm, "Nonlinear estimation of the fundamental matrix with minimal parameters," *IEEE Trans. Pattern Anal. Mach. Intell*., vol. 26, no. 3, pp. 426–432, 2004.

[9] K. Kanatani and Y. Sugaya, "Unified computation of strict maximum likelihood for geometric fitting," *J. Math. Imaging Vis.*, vol. 38, no. 1, pp. 1–13, Sep. 2010.

[10] T. Migita and T. Shakunaga, "Evaluation of epipole estimation methods with/without rank-2 constraint across algebraic/geometric error functions," in *Proc. CVPR'07. IEEE Conference on Computer Vision and Pattern Recognition*, 2007, pp. 1–7.

[11] P. H. Torr and D. W. Murray, "The development and comparison of robust methods for estimating the fundamental matrix," *Int. J. Comput. Vis.*, vol. 24, no. 3, pp. 271–300, 1997.

[12] H. S. Philip and D. W. Murray, "The development and comparison of robust methods for estimating the fundamental matrix," *International Journal of Computer Vision*, vol. 24, no. 3, 1997, 271-300.

[13] X. Armangué and J. Salvi, "Overall view regarding fundamental matrix estimation," *Image Vis. Comput*., vol. 21, no. 2, pp. 205–220, 2003.

[14] C. Harris and M. Stephens, "A combined corner and edge detector," *Dans Alvey Vision Conference*, pp. 147–151, Manchester, Royaume-Uni, janvier 1988.

[15] G. B. Yao *et al*., "Robust harris corner matching based on the quasi-homography transform and self-adaptive window for wide-baseline stereo images," *IEEE Transactions on Geoscience and Remote Sensing*, 2018.

[16] R. I. Hartley, "In defense of the eight-point algorithm," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 6, pp. 580–593, 1997.

[17] T. Xiao *et al*., "Feature matching in stereo images encouraging uniform spatial distribution," *Pattern Recognition*, vol. 48, no. 8, 2015, pp. 2530-2542.

[18] A. Chater and A. L. Asfar, "Comparaison de trois détecteurs de points d'intérêt: Harris, SIFT et SURF," 2017.

[19] A. Chater and A. L. Asfar, "Reconnaissance d'expression faciale basée sur la fusion de PCA, LBP et LPQ avec SVM," presented at Conférence Internationale CITISI'18, 2018.

[20] Y. Y. Choi *et al*., "An Efficient Fundamental Matrix Estimation for Moving Object Detection," *World Academy of Science, Engineering and Technology, International Journal of Computer and Information Engineering*, vol. 4, no. 9, 2017.

[21] R. Hartley and A. Zisserman, "Multiple view geometry in computer vision," 2004.

[22] J. H. Peter and E. M. Ronchetti, "Robust statistics, series in probability and mathematical statistics," 1981.

[23] Y. Zheng, S. Sugimoto, and M. Okutomi, "A branch and contract algorithm for globally optimal fundamental matrix estimation," in *Proc. 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011, pp. 2953–2960.

[24] L. Wang, Z. Zhang, and Z. Liu, "Efficient image features selection and weighting for fundamental matrix estimation," *IET Comput. Vis.*, vol. 10, no. 1, pp. 67–78, Feb. 2016.

[25] A. Chater and A. L. Asfar ''Detection of image descriptors and modification of the weighting function for the estimation of the fundamental matrix using robust methods," *Journal of Engineering and Applied Sciences*, 2018, vol. 13, pp. 1835-1843.

A. Chater was born on October 30, 1985. He obtained the master of physics in University Ibn Tofail kenitra in July 2015. He is a PhD student in Mohammadia Engineers School of Rabat, in engineering sciences and techniques. His research interests are in optimization and image processing methods.

A. Lasfar was born on January 10, 1971 in Salé He is higher professor at Mohammed VAgdal University, Salé Higher School of Technology, Morocco. His research focuses on compression methods, indexing by image content and image indexing and knowledge extraction from images.